(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
22 November 2001 (22.11.2001)

PCT

(10) International Publication Number
WO 01/88904 A1

(51) International Patent Classification[7]: G10L 21/02

(21) International Application Number: PCT/EP00/04601

(22) International Filing Date: 17 May 2000 (17.05.2000)

(25) Filing Language: English

(26) Publication Language: English

(71) Applicant (*for all designated States except US*): **KONINKLIJKE PHILIPS ELECTRONICS N.V.** [NL/NL]; Groenewoudseweg 1, NL-5621 BA Eindhoven (NL).

(72) Inventors; and
(75) Inventors/Applicants (*for US only*): **DEN BRINKER, Albertus, C.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL). **OOMEN, Arnoldus, W., J.** [NL/NL]; Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

(74) Agent: **GROENENDAAL, Antonius, W., M.**; Internationaal Octrooibureau B.V., Prof. Holstlaan 6, NL-5656 AA Eindhoven (NL).

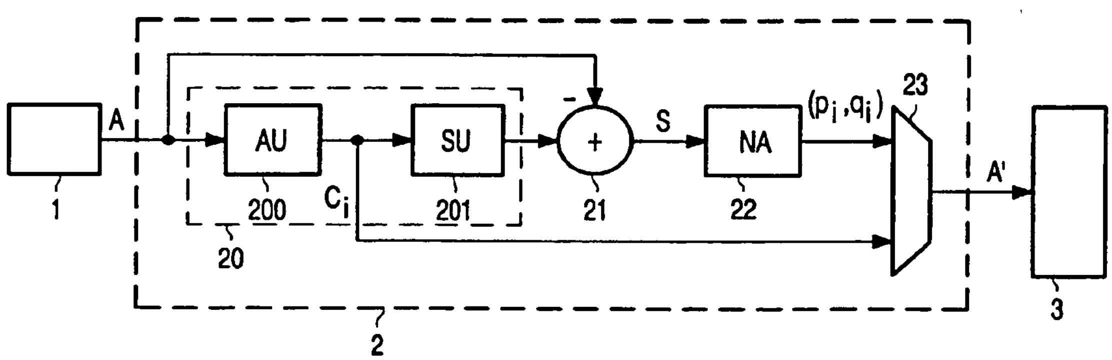
(81) Designated States (*national*): BR, CN, IN, JP, KR, MX, PL, RU, US.

(84) Designated States (*regional*): European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published:
— *with international search report*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

(54) Title: AUDIO CODING



(57) Abstract: Encoding (2) an audio signal (A) is provided, wherein basic waveforms in the audio signal (A) are determined (200), a noise component (S) is obtained (21) from the audio signal (A) by subtracting (21) the basic waveforms from the audio signal (A), a spectrum of the noise component (S) is modeled (22) by determining auto-regressive and moving-average parameters (p_i,q_i), and the auto-regressive and the moving-average parameters (p_i,q_i) are included (23) in an encoded audio signal (A') together with waveform parameters (C_i) representing the basic waveforms.

WO 01/88904 A1

Audio coding

The invention relates to audio coding.

WO 97/28527 discloses the enhancement of speech parameters by determining a background noise PSD estimate, determining noisy speech parameters, determining a noisy speech PSD estimate from the speech parameters, subtracting a background noise PSD estimate from the noisy speech PSD estimate, and estimating enhanced speech parameters from the enhanced speech PSD estimate. The enhanced parameters may be used for filtering noisy speech in order to suppress the noise or be used directly as speech parameters in speech encoding. The parameters and the PSD estimates are obtained by auto-regressive modeling. It is noted in this document that such an estimate is not a statistically consistent one, but that in speech signal processing that is not a serious problem.

An object of the invention is to provide advantageous audio coding. To this end, the invention provides a method of encoding an audio signal, a method of decoding an encoded audio signal, an audio encoder, an audio player, an audio system, an encoded audio signal and a storage medium as defined in the independent claims. Advantageous embodiments are defined in the dependent claims.

According to a first aspect of the invention, parametric ARMA modeling is used for modeling a noise component in an audio signal, which noise component is obtained by subtracting basic waveforms from the audio signal. The audio signal may comprise audio in general, like music, but also speech. ARMA modeling of the noise component according to the invention has a further advantage that for an accurate modeling of a noise component less parameters are necessary than would be the case in full AR or MA modeling with a comparable accuracy. Less parameters means, inter alia, better compression.

The invention uses an ARMA model estimation that is suitable for a real-time implementation. The invention recognizes that AR or MA models are not always sufficiently accurate or parsimonious in conveying the information of the power spectral estimate. On a logarithmic scale, with Linear Predictive Coding (LPC) methods (all-pole modeling) peaks of the function are usually well modeled but valleys are under-estimated. The reverse occurs in

an all-zero model. In audio and speech coding, a logarithmic scale is more appropriate than a linear scale. Therefore, a good fit to the power spectrum on a logarithmic scale is preferred. The model according to the invention gives a better trade-off between complexity and accuracy. The error in this model can be evaluated on a logarithmic scale.

In a first embodiment of the invention, the spectrum to be modeled is split into a first part and a second part wherein the first part is modeled by a first model to obtain auto-regressive parameters and the second part is modeled by a second model to obtain moving-average parameters. The combination of the constituent processes provides an accurate ARMA model. The splitting is preferably performed in an iterative procedure. In a method according to the invention, a non-linear optimization problem may be omitted.

In a preferred embodiment of the invention, the second modeling operation comprises the step of using the first modeling operation on a reciprocal of the second part of the target spectrum. In this embodiment, only one modeling operation needs to be defined wherein the auto-regressive parameters are obtained by modeling the first part of the spectrum and the moving-average parameters are obtained by modeling a reciprocal of the second part of the spectrum by the same, i.e. first modeling operation. Although less preferred, it is also possible to use a second modeling operation that yields moving-average parameters on the second part and, to obtain auto-regressive parameters use the same second modeling operation on a reciprocal of the first part of the spectrum.

P. Stoica and R.L. Moses, *Introduction to spectral analysis*, Prentice Hall, New Jersey, 1997, pp. 101-108, disclose parametric methods for modeling rational spectra. In general, a moving-average (MA) signal is obtained by filtering white noise with an all-zero filter. Owing to this all-zero structure, it is not possible to use an MA equation to model a spectrum with sharp peaks unless the MA order is chosen 'sufficiently large'. This is to be contrasted to the ability of the auto-regressive (AR), or all-pole, equation to model narrow-band spectra by using fairly low model orders. The MA model provides a good approximation for those spectra which are characterized by broad peaks and sharp nulls. Such spectra are encountered less frequently in applications than narrow-band spectra, so there is somewhat limited engineering interest in using MA signal model for spectral estimation. Another reason for this limited interest is that the MA parameter estimation problem is basically a non-linear one, and is significantly more difficult to solve than the AR parameter estimation problem. In any case, the types of difficulties in MA and ARMA estimation problems are quite similar.

Spectra with both sharp peaks and deep nulls cannot be modeled by either AR or MA equations of reasonably small orders. It is in these cases where the more general ARMA model, also called pole-zero model, is valuable. However, the great initial promise of ARMA spectral estimation diminishes to some extent because there is yet no well-established algorithm from both theoretical and practical standpoints for ARMA parameter estimation. The 'theoretically optimal ARMA estimators' are based on iterative procedures whose global convergence is not guaranteed. The 'practical ARMA estimators' are computational simple and often reliable, but their statistical accuracy may be poor in some cases. The prior art discloses two stage models, in which first an AR estimation is performed and thereafter an MA estimation. Both methods give inaccurate estimates or require high computational effort in those cases where the poles and zeroes of the ARMA model description are closely spaced together at positions near the unit circle. Such ARMA models, with nearly coinciding poles and zeroes of modulus close to one, correspond to narrow-band signals. In both methods, the estimation of the zeros translates to a non-linear optimization problem.

In the prior art methods according to Stoica and Moses, computational burden exists in matrix inversions. Further, it is unclear to which value the order of the AR model should be set, except that it needs to be high for zeros close to the unit circle. Therefore, the computational complexity is difficult to access. In the method according to the invention, computational burden exists in the iterative nature of the splitting process and the transformation to the frequency domain (Stoica and Moses calculate primarily in the time domain). The invention provides better results in case of zeros close to the unit circle. Furthermore, the transformation to the frequency domain opens the possibility of manipulations. An example is to make the split frequency dependent on the basis of a priori or measurement data. Another advantage is the applicability to warped frequency data, as is explained below. In order to guarantee real-time ARMA modeling, a fast transformation to the frequency domain should be applied, e.g. Welch's averaged periodogram method which is well known in the art.

Auto-regressive and moving average parameters can be represented in different ways by e.g. polynomials, zeros of the polynomials (together with a gain factor), reflection coefficients or log(Area) ratios. In an audio coding application, representation of the auto-regressive and moving average parameters is preferably in log(Area) ratios. The auto-regressive and moving average parameters that are determined in the ARMA modeling according to the invention are combined to obtain the filter parameters that are transmitted.

US-A 5,943,429 discloses a spectral subtraction noise suppression method in a frame based digital communication system. The method is performed by a spectral subtraction function which is based on an estimate of the power spectral density of . background noise of non-speech frames and an estimate of the power spectral density of speech frames. Each speech frame is approximated by a parametric model that reduces the number of degrees of freedom. The estimate of the power spectral density of each speech frame is estimated from the approximative parametric model. Also in this case, the parametric model is an AR model.

US-A 4,188,667 discloses an ARMA filter and a method for obtaining the parameters for such filter. The first step of this method involves performing an inverse discrete Fourier transform of the arbitrary selected frequency spectrum of amplitude to obtain a truncated sequence of coefficients of a stable pure moving-average filter model, i.e. the parameters of a non-recursive filter model. The truncated sequence of coefficients, which has N+1 terms, is then convolved with a random sequence to obtain an output associated with the random sequence. A time-domain, convergent parameter identification is then performed, in a manner that minimizes an integral error function norm, to obtain the near minimum order auto-regressive and moving-average parameters of the model having the desired amplitude- and phase-frequency responses. The parameters are identified off-line. The object of this embodiment is to provide a minimum or near minimum stable ARMA filter. The parameters are determined in a batch filter program.

In general, estimating a power spectral density function differs from characterizing a linear system in that, inter alia, in such characterization, the input and output signals are available and used, whereas in estimating a power spectral density function, only the power spectral density function is available (not an associated input signal).

The aforementioned and other aspects of the invention will be apparent from and elucidated with reference to the embodiments described hereinafter.

In the drawings:

Fig. 1 shows an illustrative embodiment comprising an audio encoder according to the invention;

Fig. 2 shows an illustrative embodiment comprising an audio player according to the invention;

Fig. 3 shows an illustrative embodiment of an audio system according to the invention; and

Fig. 4 shows an exemplary mapping function m.

The drawings only show those elements that are necessary to understand the invention.

The invention is preferably applied in audio and speech coding schemes in which synthetic noise generation is employed. Typically, the audio signal is coded on a frame to frame basis. The power spectral density function (or a possibly non-uniform sampled version thereof) of the noise in a frame is estimated and a best approximation of the function from a set of squared amplitude responses of a certain class of filters is found. In one embodiment of the invention, an iterative procedure is used to estimate an ARMA model based on existing low-complexity techniques for fitting AR and MA models to the power spectral density function.

Fig. 1 shows an exemplary audio encoder 2 according to the invention. An audio signal A is obtained from an audio source 1, such as a microphone, a storage medium, a network etc. The audio signal A is input to the audio encoder 2. The audio signal A is parametrically modeled in the audio encoder 2 on a frame to frame basis. A coding unit 20 comprises an analysis unit (AU) 200 and a synthesis unit (SU) 201. The AU 200 performs an analysis of the audio signal and determines basic waveforms in the audio signal A. Further, the AU 200 produces waveform parameters or coefficients C_i to represent the basic waveforms. The waveform parameters C_i are furnished to the SU 201 to obtain a reconstructed audio signal, which consists of synthesized basic waveforms. This reconstructed audio signal is furnished to a subtractor 21 to be subtracted from the original audio signal A. This rest signal S is regarded as a noise component of the audio signal A. In a preferred embodiment, the coding unit 20 comprises two stages: one that performs transient modeling, and another that performs sinusoidal modeling on the audio signal after subtraction of the modeled transient components.

According to an aspect of the invention, the power spectral density function of the noise component S in the audio signal A is ARMA modeled resulting in auto-regressive parameters p_i and moving-average parameters q_i. The spectrum of the noise component S is modeled according to the invention in noise analyzer (NA) 22 to obtain filter parameters (p_i,q_i). The estimation of the parameters (p_i,q_i) is performed by determining filter parameters of a filter in NA 22 which has a transfer function H^{-1} that makes the function S after filtering, i.e. $H^{-1}(S)$, spectrally as flat as possible, i.e. 'whitening the frequency spectrum'. In a decoder, a reconstructed noise component can be generated which has approximately the same

properties as the noise component S by filtering white noise with a filter with transfer function H that is opposite to the filter used in the encoder. The filtering operation of this opposite filter is determined by the ARMA parameters p_i and q_i. The filter parameters (p_i,q_i) are included together with the waveform parameters C_i in an encoded audio signal A' in a multiplexer 23. The audio stream A' is furnished from the audio encoder to an audio player over a communication channel 3, which may be a wireless connection, a data bus or a storage medium, etc.

An embodiment comprising an audio player 4 according to the invention is shown in Fig. 2. An audio signal A' is obtained from the communication channel 3 and de-multiplexed in de-multiplexer 40 to obtain the parameters (p_i,q_i) and the waveform parameters C_i that are included in the encoded audio signal A'. The parameters (p_i,q_i) are furnished to a noise synthesizer (NS) 41. The NS 41 is mainly a filter with a transfer function H. A white noise signal y is input to the NS 41. The filtering operation of the NS 41 is determined by the ARMA parameters (p_i,q_i). By filtering the white noise y with the NS 41, that is opposite to the filter (NA) 22 used in the encoder 2, a noise component S' is generated which has approximately the same stochastic properties as the noise component S in the original audio signal A. The noise component S' is added in adder 43 to other reconstructed components, which are e.g. obtained from a synthesis unit (SU) 42 to obtain a reconstructed audio signal (A''). The SU 42 is similar to the SU 201. The reconstructed audio signal A'' is furnished to an output 5, which may be a loudspeaker, etc.

Fig. 3 shows an audio system according to the invention comprising an audio encoder 2 as shown in Fig. 1 and an audio player 4 as shown in Fig. 2. Such a system offers playing and recording features. The communication channel 3 may be part of the audio system, but will often be outside the audio system. In case the communication channel 3 is a storage medium, the storage medium may be fixed in the system or be a removable disc, memory stick, tape etc.

Below, the modeling of the spectrum of S is further described. Suppose S is a power spectral density function of a discrete-time real valued signal. Further, S is a real-valued function defined on the interval $I = (-\pi,\pi)$. S is assumed to be symmetric with min $(S) > 0$ and max $(S) < \infty$. For convenience, it is assumed that the logarithmic mean of S equals zero, i.e.

$$\frac{1}{2\pi}\int_I \ln S(\theta)d\theta = 0 \qquad (1)$$

The extension to cases with a mean on the log scale unequal to zero is straight forward, but can be handled in various ways. Note that S can be derived from samples of an actually measured power spectral density function by suitable interpolation and normalization.

Let H be a rational transfer function according to $H = B / A$ with $A = \prod_{i=1}^{N}(1 - z^{-1}p_i)$ and $B = \prod_{i=1}^{M}(1 - z^{-1}q_i)$. Here, p_i and q_i are the poles and the zeros of the transfer function H, respectively. Note, that the logarithmic mean of $|H|^2$ also equals zero.

The target function is approximated by the squared modulus of H, i.e. $S \approx |H|^2$.

A measure for the correctness of the approximation is introduced by:

$$J = \frac{1}{2\pi}\int\frac{1}{2}(\ln S - \ln|H|^2)^2\,d\theta \tag{2}$$

The criterion (2) can be rewritten to

$$J = \frac{1}{2\pi}\int \ln(S/|H|^2) + \frac{1}{2}(\ln(S/|H|^2))^2\,d\theta \tag{3}$$

in view of the fact that both S and $|H|^2$ have a logarithmic mean equal to zero. If furthermore, $S(\theta)/|H(e^{j\nu}|^2 \approx 1$ for each θ, the criterion (2) is approximated by J'-1, where

$$J' = \frac{1}{2\pi}\int\frac{S}{|H|^2}\,d\theta \tag{4}$$

This means that in the neighborhood of the optimal solution, the criteria (2) and (4) are practically equal.

It is well known that in the case that $H = 1/A$ (i.e. $B = 1$), (4) is associated with Forward Linear Prediction (FLP), which is an example of an LPC method. Therefore, the polynomial A can be found by calculating (or at least approximating) the auto-correlation function associated with S and solving the Wiener-Hopf equations. The qualitative results of such a procedure are also well known. The above sketched procedure will give good approximations to the peaks of S (when measured or visualized on a logarithmic scale) but usually provides only poor fits to the valleys of S. To conclude the above, a standard procedure is available for estimating an all-pole model from the power spectral density function, which provides an approximation to the optimal solution with (2) and which basically is good at modeling the peaks of S.

It is noted that peaks and valleys of $\ln S$ have essentially the same characteristic except for a reversal of sign: a peak is a positive excursion, whereas a trough is

a negative one. Consequently, taking $\hat{S} = 1/S$, an all-zero model can be estimated by using the above sketched procedure for an all-pole model. From the result of this procedure, a good fit to the valleys of S is expected, but only poor or at most fair fits to the peaks of S.

An object of the invention is to provide a good representation of S for both the peaks and the valleys. In an embodiment of the invention, an ARMA model is provided in which all-pole modeling and all-zero modeling are combined in the following way. S is split in two parts as $S = S_A / S_B$. From S_A an all pole model is estimated yielding the polynomial A and from S_B an all-zero model is estimated yielding the polynomial B. The combination $|H|^2 = |B|^2 / |A|^2$ is considered an approximation of S.

According to a preferred aspect of the invention the split of S is performed in an iterative process. The iteration step is called l. At each step of the iteration, a new split $S_{A,l}$ and $S_{B,l}$ is generated and the corresponding estimates A_l and B_l are calculated. A given subdivision of S in S_A and S_B is used to start with and thereafter parts of S_B that are not modeled accurately are attributed to S_A and vice versa. At step l-1 in the iterative scheme, $H_{l-1} = B_{l-1} / A_{l-1}$. Hereafter, the partial functions $S_{A,l} = S / |B_{l-1}|^2$ and $S_{B,l} = 1/S|A_{l-1}|^2$ are considered. In this way, from S those parts that can be modeled accurately by the all-pole model are excluded from contributing to S_B. Similarly, those parts of S that could be modeled by an all-zero filter are excluded from S_A. From $S_{A,l}$ and $S_{B,l}$ the functions A_l and B_l are estimated. In this way, parts which in the previous iteration could not be modeled appropriately are swapped.

For a next step, preferably, the following four possible combinations are considered:

$$G_0 = B_{l-1} / A_{l-1} \qquad G_1 = B_{l-1} / A_l$$

$$G_2 = B_l / A_{l-1} \qquad G_3 = B_l / A_l$$

The best fit to S of these four candidate filters is defined as the one with minimum error; the associated filter is the final result of step l. Preferably, H_l (and thus A_l and B_l) is selected as the best of the candidates G_i with i = 0,1,2,3 on a logarithmic criterion according to

$$H_l = \arg\min_{G_i} \frac{1}{2\pi} \int_I (\ln S - \ln|G_i|^2)^2 \, d\theta \qquad (5)$$

From here, the procedure is proceeded with step $l + 1$, by taking $S_{A,l+1} = S / |B_l|^2$ and $S_{B,l+1} = 1/S|A_l|^2$.

Any common stop procedure can be used, e.g. a maximum number of iterations, a sufficient accuracy of the current estimate, or insufficient progress in going from one step to another.

A slightly different procedure performs the AR and MA modeling alternately. If the previous step returned a refined estimate of the numerator B_{l-1}, then

$$S_{A,l} = S / |B_{l-1}|^2$$

and calculate A_l. B_l is taken as B_{l-1}.

If the previous step returned a refined estimate of the numerator A_{l-1}, then

$$S_{B,l} = 1 / S|A_{l-1}|^2$$

and calculate B_l. A_l is taken as A_{l-1}.

From A_l and B_l, H_l is constructed and the error evaluated (e.g. a mean squared difference on a log scale)

There are many alternatives to initialize the iterative scheme. Without limitation, the following possibilities are mentioned:

First, a simple way of initializing is provided by taking $S_{A,0} = S$ and $S_{B,0} = 1$ and $S_{A,0} = 1$ and $1/S_{B,0} = S$. Next, A_0 and B_0 are calculated. From these two initial estimates, a best fit (according to some criterion) is chosen. In this way, the first guess is either an all-pole or an all-zero model.

Second, S may be split in equal parts according to $S_{A,0} = 1/S_{B,0} = \sqrt{S}$.

Third, since S_A should contain the peaks and S_B the valleys, a favorable split is to attribute everything above a mean logarithmic level (e.g. above zero) to $S_{A,0}$ and anything below said level to $S_{B,0}$. This division may be made at the global logarithmic mean, but also at some local logarithmic mean.

Fourth, a further splitting process takes into account that in power spectral density functions on a logarithmic scale, poles and zeros close to the unit circle give rise to pronounced peaks and valleys, respectively. The data S is split on the notion that peaks and valleys in $\log S$ are more appropriately handled by the all-pole and all-zero model, respectively. Define:

$P = \log S$

$P_A = \log S_A$

$P_B = \log S_B$

Consider the mapping function m with $m : \Re \rightarrow [-1,1]$. The mapping function will typically be a non-decreasing, point-symmetric sigmoidal function in view of the symmetry of pole

and zero behavior on a log scale. However, non-symmetric functions can be used as well and have the effect of giving more weight to either the pole or the zero modeling. An exemplary mapping function m is shown in Fig. 4.

Consider the following initial split:

$$P_A = \frac{1+m(P)}{2} P$$

$$P_B = -\frac{1-m(P)}{2} P$$

In this way, positive excursion (peaks) of P are pre-dominantly attributed to P_A and, consequently, modeled by the all-pole filter. Negative excursions (valleys) of P are mostly attributed to P_B and, consequently, modeled by the all-zero filter. From P_A and P_B, S_A and S_B are constructed and, next A_0 and B_0 are calculated.

There are two limiting cases of m (which are similar to the second and the third initialization as discussed above):

- m = 0, then $S_{A,0} = 1/S_{B,0} = \sqrt{S}$
- m is a signum function: $m(x) = \begin{cases} -1, x<0 \\ 0, x=0 \\ 1, x>0 \end{cases}$

In this case:

$$S_A(x) = \begin{cases} S(x), S(x)>1 \\ 1, S(x)\le 1 \end{cases}$$

$$1/S_B(x) = \begin{cases} S(x), S(x)<1 \\ 1, S(x)\ge 1 \end{cases}$$

The proposed spectrum modeling is very apt at modeling peaks and valleys since, basically, these constitute the patterns generated by the degrees of freedom offered by the poles and zeros. Consequently, the procedure is sensitive to outliers: rather than smoothing, these will appear in the approximation. Therefore, the input data S has to be an accurate estimate (in the sense of a small ratio of standard deviation and mean per frequency sample) or S must be pre-processed (e.g. smoothed) in order to suppress undesired modeling of outliers. This observation holds especially if the number of degrees of freedom in the model is relatively large with respect to the number of data points on which the power spectral density function is based.

Convergence can not be established without knowledge of the actual optimization steps A and B and the selection criterion. It is not guaranteed that the error reduces at every step in the iteration process.

In many cases, it is desired to have a good approximation of the power spectral density function on a logarithmic scaled frequency axis. For example, it is common practice to evaluate the result of a fit on a spectrum visually in the form of a Bode plot. Similarly, for audio and speech applications, the preferred scale would be a Bark or Equivalent Rectangular Bandwidth (ERB) scale which is more or less a logarithmic scale. The method according to the invention is suitable for frequency-warped modeling. The spectral density measurements can be calculated on any frequency grid whatsoever. Under the condition that the frequency warping is close to that of a first-order all-pass section, this can be re-wrapped while maintaining the order of the ARMA model.

It should be noted that the above-mentioned embodiments illustrate rather than limit the invention, and that those skilled in the art will be able to design many alternative embodiments without departing from the scope of the appended claims. In the claims, any reference signs placed between parentheses shall not be construed as limiting the claim. The word 'comprising' does not exclude the presence of other elements or steps than those listed in a claim. The invention can be implemented by means of hardware comprising several distinct elements, and by means of a suitably programmed computer. In a device claim enumerating several means, several of these means can be embodied by one and the same item of hardware. The mere fact that certain measures are recited in mutually different dependent claims does not indicate that a combination of these measures cannot be used to advantage.

In summary, encoding an audio signal is provided, wherein basic waveforms in the audio signal are determined, a noise component is obtained from the audio signal by subtracting the basic waveforms from the audio signal, a spectrum of the noise component is modeled by determining auto-regressive and moving-average parameters, and the auto-regressive and the moving-average parameters are included in an encoded audio signal together with waveform parameters representing the basic waveforms.

CLAIMS:

1. A method of encoding (2) an audio signal (A), comprising the steps of:

determining (200) basic waveforms in the audio signal (A);

obtaining (21) a noise component (S) from the audio signal (A) by subtracting (21) the basic waveforms from the audio signal (A);

modeling (22) a spectrum of the noise component (S) by determining auto-regressive and moving-average parameters (p_i,q_i); and

including (23) the auto-regressive and the moving-average parameters (p_i,q_i), and waveform parameters (C_i) representing the basic waveforms in an encoded audio signal (A').

2. A method of decoding (4) an encoded audio signal (A'), comprising the steps of:

receiving (40) an encoded audio signal (A') comprising waveform parameters (C_i) representing basic waveforms and auto-regressive and moving-average parameters (p_i,q_i) representing a spectrum of a remaining noise component;

filtering (41) a white noise signal (y) to obtain a reconstructed noise component (S'), which filtering is determined by the auto-regressive parameters (p_i) and the moving-average parameters (q_i);

synthesizing (42) basic waveforms based on the waveform parameters (C_i); and

adding (43) the reconstructed noise component (S') to the synthesized basic waveforms to obtain a decoded audio signal (A'').

3. An audio encoder (2) comprising:

means (200) for determining basic waveforms in the audio signal (A);

means for (21) obtaining a noise component (S) from the audio signal (A) by subtracting (21) the basic waveforms from the audio signal (A);

means (22) for modeling a spectrum of the noise component (S) by determining auto-regressive and moving-average parameters (p_i,q_i); and

means (23) for including the auto-regressive and the moving-average parameters (p_i,q_i), and waveform parameters (C_i) representing the basic waveforms in an encoded audio signal (A').

4. An audio player (4) comprising:

means (40) for receiving an encoded audio signal (A') comprising waveform parameters (C_i) representing basic waveforms and auto-regressive and moving-average parameters (p_i,q_i) representing a spectrum of a noise component;

means (41) for filtering a white noise signal (y) to obtain a reconstructed noise component (S'), which filtering is determined by the auto-regressive parameters (p_i) and the moving-average parameters (q_i);

means (42) for synthesizing basic waveforms based on the waveform parameters (C_i); and

means (43) for adding the reconstructed noise component (S') to the synthesized basic waveforms to obtain a decoded audio signal (A'').

5. An audio system comprising an audio encoder (2) as claimed in claim 3 and an audio player (4) as claimed in claim 4.

6. An encoded audio signal (A') comprising:

waveform parameters (C_i) representing basic waveforms;

auto-regressive parameters and moving-average parameters (p_i,q_i) representing a spectrum of a remaining noise component (S).

7. A storage medium (3) on which an encoded audio signal (A') as claimed in claim 6 is stored.

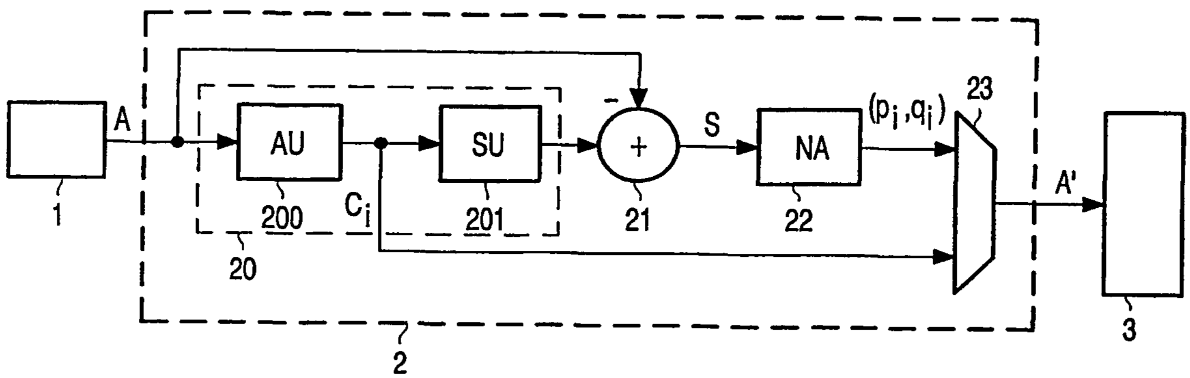


FIG. 1

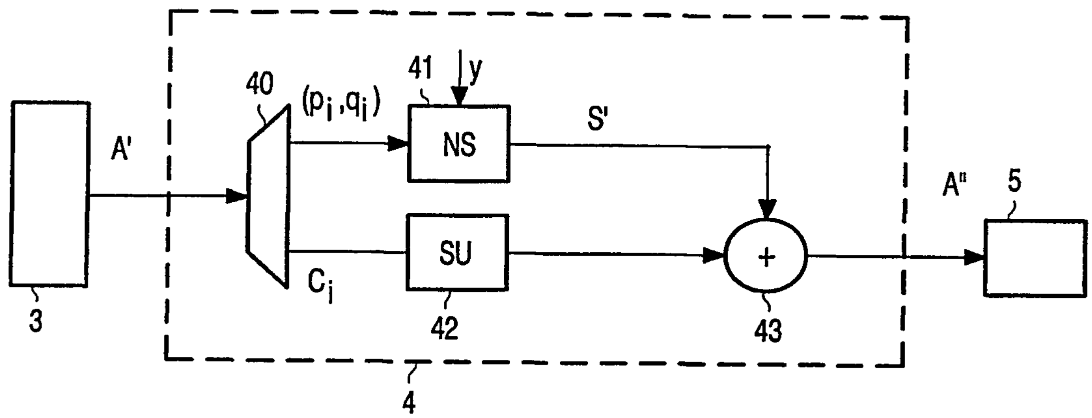


FIG. 2

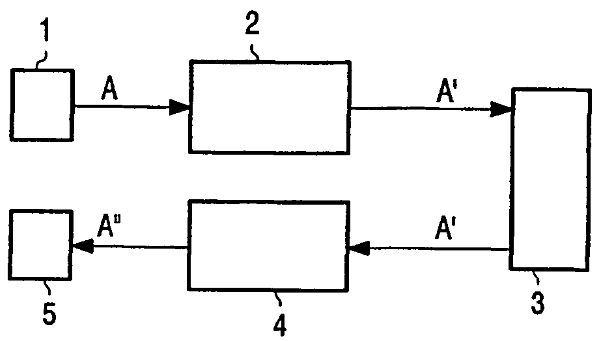


FIG. 3

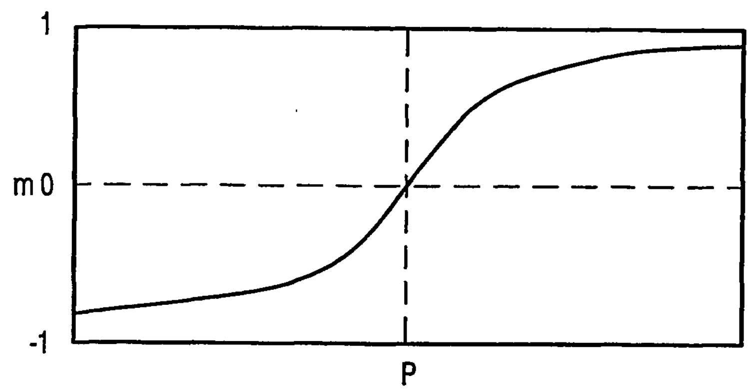


FIG. 4

P ENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference NL000288WO	**FOR FURTHER ACTION** see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.	
International application No. PCT/EP 00/ 04601	International filing date *(day/month/year)* 17/05/2000	(Earliest) Priority Date *(day/month/year)*
Applicant KONINKLIJKE PHILIPS ELECTRONICS N.V.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 2 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. **Basis of the report**
   a. With regard to the **language,** the international search was carried out on the basis of the international application in the language in which it was filed, unless otherwise indicated under this item.
      [ ] the international search was carried out on the basis of a translation of the international application furnished to this Authority (Rule 23.1(b)).
   b. With regard to any **nucleotide and/or amino acid sequence** disclosed in the international application, the international search was carried out on the basis of the sequence listing :
      [ ] contained in the international application in written form.
      [ ] filed together with the international application in computer readable form.
      [ ] furnished subsequently to this Authority in written form.
      [ ] furnished subsequently to this Authority in computer readble form.
      [ ] the statement that the subsequently furnished written sequence listing does not go beyond the disclosure in the international application as filed has been furnished.
      [ ] the statement that the information recorded in computer readable form is identical to the written sequence listing has been furnished

2. [ ] **Certain claims were found unsearchable** (See Box I).
3. [ ] **Unity of invention is lacking** (see Box II).

4. With regard to the **title,**
   [X] the text is approved as submitted by the applicant.
   [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract,**
   [X] the text is approved as submitted by the applicant.
   [ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this international search report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is Figure No. 1
   [X] as suggested by the applicant.
   [ ] because the applicant failed to suggest a figure.
   [ ] because this figure better characterizes the invention.
   [ ] None of the figures.

INTERNATIONAL SEARCH REPORT

International Application No
PCT/EP 00/04601

A. CLASSIFICATION OF SUBJECT MATTER
IPC 7 G10L21/02

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 7 G10L

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

EPO-Internal, INSPEC, WPI Data, PAJ, IBM-TDB, COMPENDEX

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
A	US 5 721 694 A (GRAUPE DANIEL) 24 February 1998 (1998-02-24) abstract; claim 1	1-4
A	US 4 188 667 A (BEEX ALOYSIUS A ET AL) 12 February 1980 (1980-02-12) cited in the application abstract	1-4
A	US 5 943 429 A (HAENDEL PETER) 24 August 1999 (1999-08-24) cited in the application abstract; claim 1	1-4
A	US 5 717 724 A (SATO TOMONORI ET AL) 10 February 1998 (1998-02-10) abstract column 9, line 59 -column 10, line 15	1-4

[ ] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
25 January 2001	01/02/2001

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL – 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Van Doremalen, J

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/EP 00/04601

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 5721694 A	24-02-1998	NONE	
US 4188667 A	12-02-1980	JP 1355090 C	24-12-1986
		JP 52125251 A	20-10-1977
		JP 61018887 B	14-05-1986
		DE 2707607 A	01-09-1977
		DK 76877 A	24-08-1977
		NL 7701950 A	25-08-1977
US 5943429 A	24-08-1999	SE 505156 C	07-07-1997
		AU 696152 B	03-09-1998
		AU 4636996 A	21-08-1996
		BR 9606860 A	25-11-1997
		CA 2210490 A	08-08-1996
		DE 69606978 D	13-04-2000
		DE 69606978 T	20-07-2000
		EP 0807305 A	19-11-1997
		ES 2145429 T	01-07-2000
		FI 973142 A	30-09-1997
		JP 10513273 T	15-12-1998
		SE 9500321 A	31-07-1996
		WO 9624128 A	08-08-1996
US 5717724 A	10-02-1998	JP 8130513 A	21-05-1996